The Impact of Carbon Emission on Global Warming Based on LSTM Model and PCA

Chenyi Zhu*

School of Mathematics and Statistics, Hunan Normal University, Changsha, China

*Corresponding author: zhuchenyi@hunnu.edu.cn

Abstract. Climate related topic has become a big topic in recent years. Frequent occurrence of extreme phenomenon has made people to are about the climate issue throughout the world, especially about the global warming. This study uses LSTM model to predict the temperature anomaly in the next 100 years and makes PCA to analyze the main contributor to global warming. The temperature anomaly this article used as the dependent variable will stay in a relatively stable stage and the highest anomaly is 1.24℃ while the lowest is 0.25℃. Among carbon dioxide emission, forest fires and volcanic eruption, the main factor that contributes to global warming id CO_2. At end of the article, some practical suggestions are put forward to reduce carbon emission in normal life. This article aims to present a basic knowledge of global warming and climate change, and provide people with some measures to implement low carbon.

Keywords: Carbon emission; global warming; LSTM; PCA.

1. Introduction

Since the industrial revolution in the early 20th century, global climate change has always been a big topic in the world. One of the main reasons contributing to climate change is the intensification of the greenhouse effect [1]. According to the report published by NOAA, the level of carbon dioxide (CO_2) in 2022 was the highest level since monitoring [2]. Human activities, such as burning fossil fuels for energy is one of the contributors to the increase of carbon dioxide concentration [3]. Based on IPCC's sixth assessment report, extreme weather phenomenon is becoming widespread, which means the earth is warming day after day [4]. To slow down the warming tendency, or stop the tendency, one has to reduce the emissions of carbon by human activities [5] so that one can reduce the amount of carbon in the atmosphere [6].

Scholars have already started some research about climate change, especially ways to analyze and predict the carbon amount in the atmosphere. Based on their study, many practical methods are put forward aiming to reduce the carbon emission. Deng et al. quantitatively simulated the relationship between carbon emissions and the rise of temperature under the condition of non-uniform distribution of global carbon dioxide concentration by using spatial autocorrelation analysis and spatial simultaneous equations model. According to their research, carbon dioxide concentrations in the northern and southern hemispheres showed different patterns. It is necessary to take the difference into consideration when calculating the relationship [7]. In addition, Gao et al. made a novel fractional grey Riccati model to predict the emission of carbon in the US, China and Japan. The model can work very well in short-term prediction. The results show that three countries would reduce their carbon emission in 2025 under current policies [8]. Since China is the largest carbon emitter, Huang et al. made a prediction about the carbon emission in China based on long short-term memory (LSTM) and provided some solutions to reduce carbon emission [9]. Moreover, Bernardina Algieri, Leonardo Iania and Arturo Leccadito made a comparison about the accuracy among 13 different prediction models. Their study showed that drought and Antarctic wind indicators contribute a lot to the short-term and medium-term prediction of carbon emission [10]. According to Liu and Wang's research on Beijing, China, expanding employment, improving public transportation, and decreasing the use of household electricity and waste would help to reduce carbon emission in the air [11].

It is very meaningful to predict future carbon emission and analyze the contributors that lead to the increase of carbon emission. Based on the results, the government can implement some effective

measures to reduce carbon emission. The main body of this article is separated into 3 parts: data and method, simulation results and analysis, as well as limitations and future outlooks. In data and method, this article will explain the sources of the data and introduce the models used, including data processing and evaluation metrics. In simulation results and analysis, the results predicted by the models will be presented. The results will include the main factors that cause the warming and tell the reader the changing trend and increment of carbon emissions in the next few years. This study will also discuss and analyze these results in this part. Finally, this article will analyze the stability of the models used in the report and their limitations of the models used in this article. According to the analysis, some methods to reduce carbon emission will be put forward for reference.

2. Data and Method

2.1. Model

Considering that the change of CO_2 and temperature anomaly is nonlinear and unstable, this article decided to choose both the long short-term memory neural network (LSTM) and autoregressive integrated moving average (ARIMA) to make the prediction of global temperature in about the next 50 and 100 years [12]. Additionally, this article also uses Principal component analysis (PCA) to find the main contributors of the temperature change and then provides some useful suggestions based on the results. Long short-term memory neural network is a useful model to solve long-term time series problems, especially for long term prediction. The most important part of the model is its 'Memory Block'. The main structure of a single LSTM is shown in Fig. 1. After running the model, loss calculation needs to be done. This article uses mean square error, or known as MSE and Adam to calculate loss. The function is given as:

$$loss(\hat{y}, y) = \frac{1}{n}\sum(\hat{y}_i - y_i)^2 \tag{1}$$

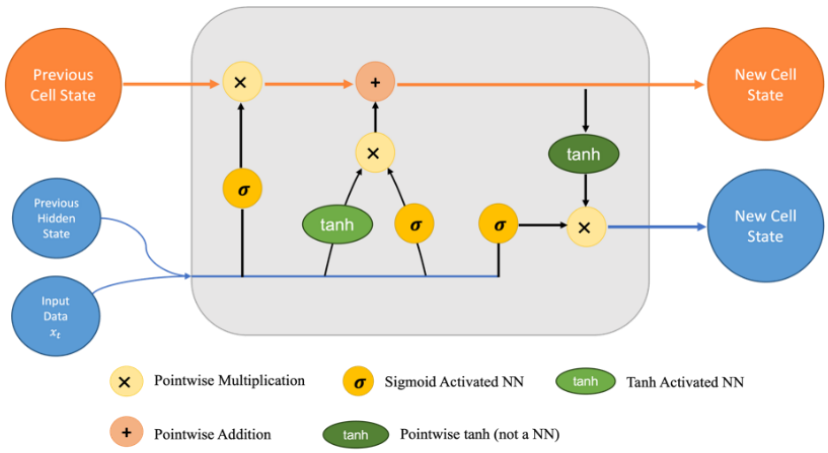


Fig. 1 Internal structure of single LSTM.

PCA is often used to determine which factor has the most important effect on the global temperature. The main idea of PCA is to decrease the dimension of the data. To meet this goal, the covariance matrix of the data and eigenvalues of the covariance matrix should be calculated. Then choose the matrix corresponding to the largest eigenvalue, and this matrix is the main point of the result. Assuming that $\vec{v}$ is an eigenvector of matrix A, λ then the relationship between the vector and the matrix is

$$A\vec{v} = \lambda\vec{v} \tag{2}$$

Then orthogonalize and unitize all the eigenvectors of matrix A to obtain a set of orthogonal unit vectors, so that the matrix can be decomposed into eigenvalues. Then the matrix A can be written as

$$A = Q\sum Q^{-1} \tag{3}$$

Here, Q is a matrix composed of the eigenvectors of A, it is a diagonal matrix, and elements on the diagonal are the eigenvalues of A. Given the data set $X = \{x_1, x_2, \cdots, x_n\}$, and k is the final dimension. The covariance matrix of it is given by

$$A = \frac{1}{n} X X^{-1} \tag{4}$$

Then one solves the equation to get its eigenvalues and eigenvectors. Arrange the eigenvalues from the largest to the smallest and pick up the first k ones. Let the matrix P made up from eigenvectors corresponding to the eigenvalues that have just selected. The final matrix is given by

$$Y = PX \tag{5}$$

2.2. Data Source

In this article, the climate data, which are annual average temperature anomaly and carbon dioxide emission, are from NOAA NCEI. When doing climate studies, temperature anomaly is more often used than absolute temperature [13]. In this article, the baseline of the temperature anomaly is the average temperature of 1901 to 2000, which is 13.9°C. The data used for discover the relationship between temperature anomaly and carbon dioxide emission is selected from 1980 to 2022 to make sure they have the same time scale. The data used for temperature prediction is monthly average temperature anomaly from January 1850 to July 2023 to make sure the LSTM model can perform well. Since the LSTM model is often used to deal with problems with very large dataset. Among many factors that contribute to the change of temperature, volcanic eruptions, forest fires and CO_2 emission have the most significant impact on the climate The time scales are from 2001 to 2021. This article uses the loss of forest are caused by fires in replacement of forest fires. Before doing PCA, a correlation analysis of these factors should be done.

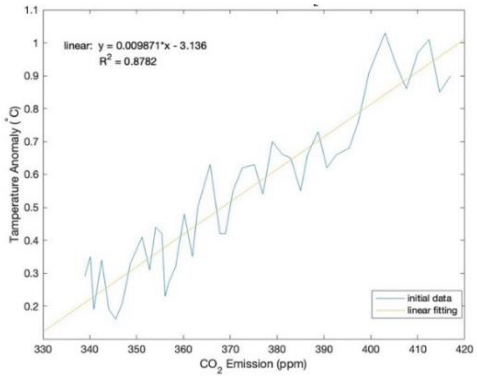


Fig. 2 Temperature Anomaly versus Carbon Dioxide Emission.

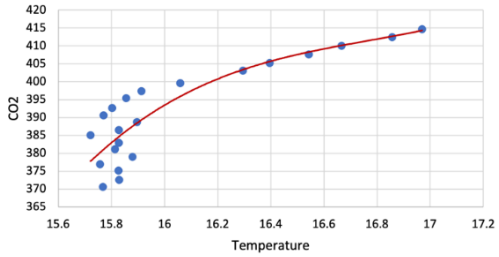


Fig. 3 Scatter plot: CO_2 and temperature.

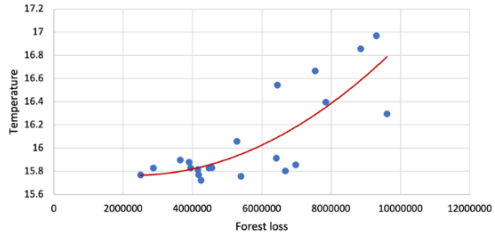


Fig. 4 Scatter plot: Forest loss and temperature.

3. Results and Discussion

3.1. Data Processing and Feature Engineering

Seen from Fig. 2 the annual global temperature anomaly is increasing with the increase of CO2 emission generally. It is apparent that CO2 emission has a very significant influence on global temperature. In this case, the independent variable is CO2 emission, and the dependent variable is temperature anomaly. In the LSTM model, the data are divided into two parts. This article uses data from 1850 to 2000 for training the model, and data from 2001 to 2022 for prediction. RMSE and loss are used to measure the performance of the model. RMSE is root-mean-squared error. It optimal in Gaussian errors [14], in which case the observed value and predicted value are independent with each other. As shown in Fig. 3, Fig. 4 and Fig. 5, the relationship between CO_2 and temperature, forest loss and temperature, volcanic eruption and temperature are showing either a increasing trend or a decreasing trend. In this case, Spearman's correlation coefficient is suitable for correlation analysis. By using SPSS, the relationship between CO2 and temperature is very significant at the significant value of 0.01. CO2 emission has a strong relationship to temperature. So as the relationship between the other two factors and temperature, respectively.

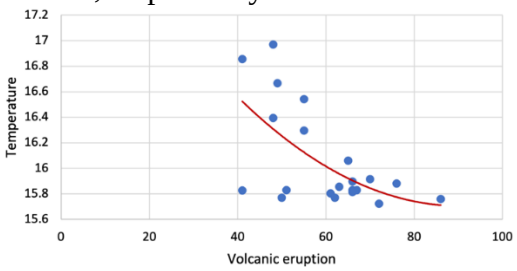


Fig. 5 Scatter plot: Volcanic eruption and temperature.

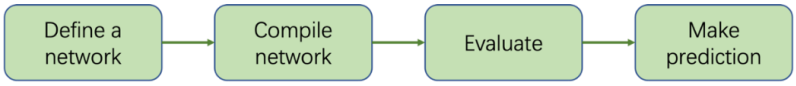


Fig. 6 Flow chart of the solution of LSTM.

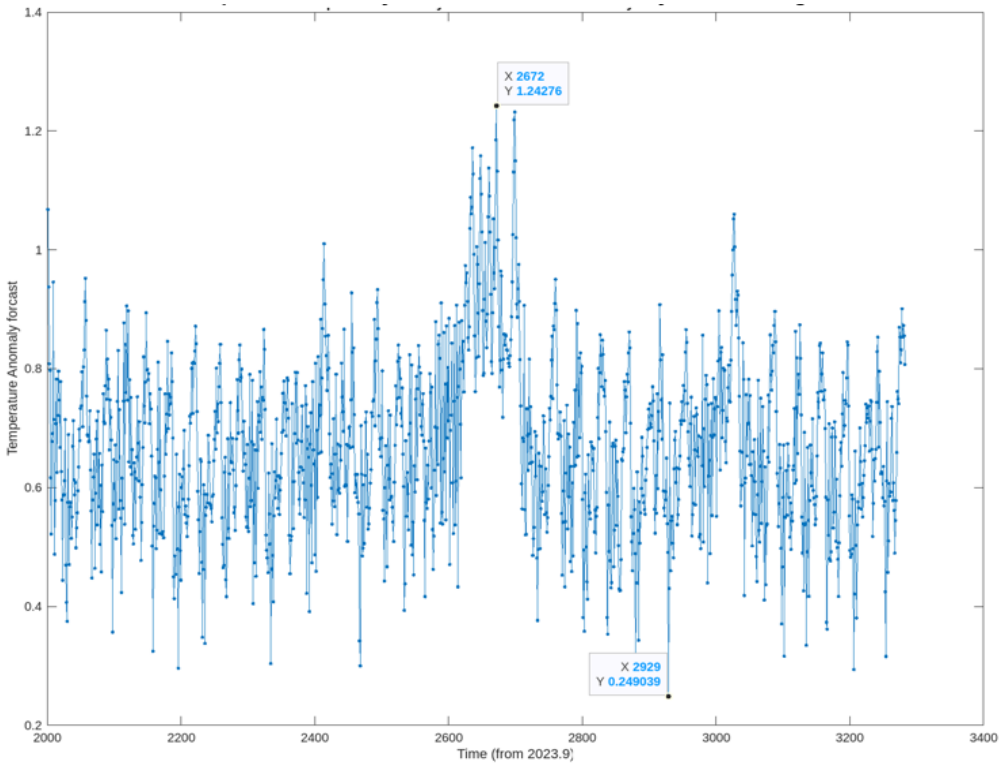


Fig. 7 Monthly Temperature Anomaly Prediction for the next 100 years from July 2023

3.2. Model Performances

The whole solving process can be divided into four steps, as is shown in Fig. 6. In this article, the number of iterations of the model is 1000. The prediction result shows that the average global temperature anomaly will be between 0.25℃ and 1.24℃, see in Fig. 7. Compare to the general trend before the forecast time, the temperature change would slow down for some reasons (seen from Fig. 8). The RMSE and loss with the number of iterations finally turn to under 0.1, which means the model performs well in both training and predicting (seen from Fig. 9). However, for LSTM, the numerical results can be easily affected by the random initial value of the neural network, so it is possible that the results of prediction will not be stable enough. In PCA, the final Among all the possible factors, there is only one factor that contributes the most to global warming: CO_2, which is very comprehensible.

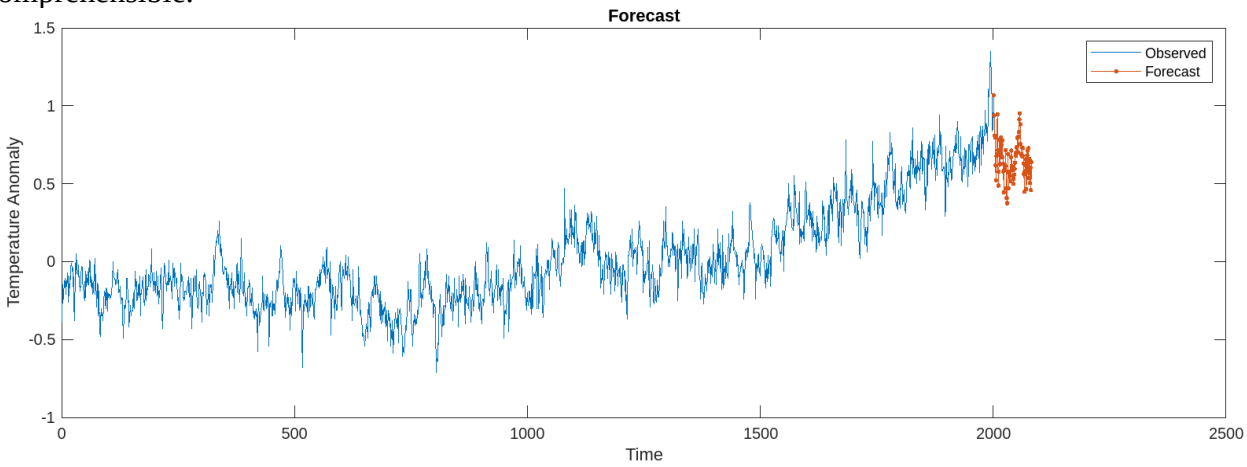


Fig. 8 The observed data curve and forecast curve.

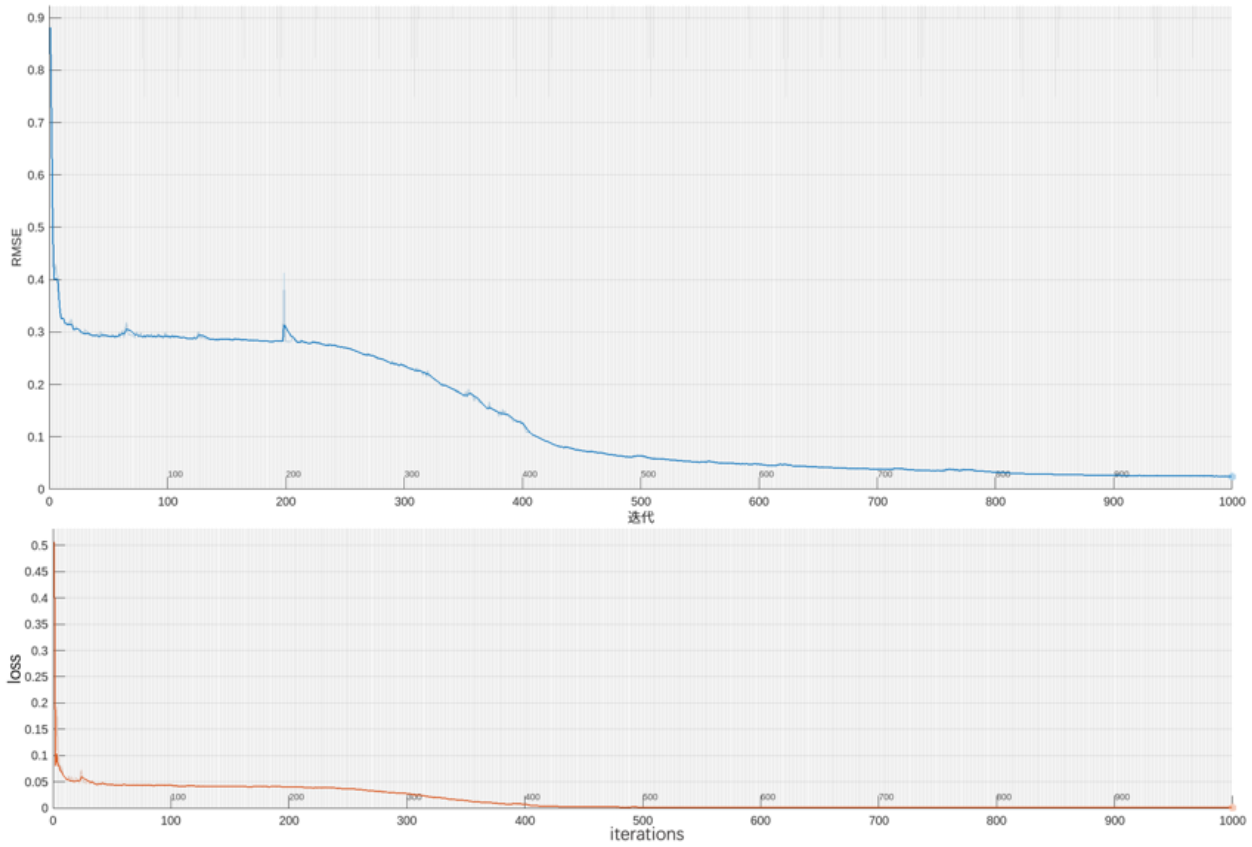


Fig. 9 The RMSE and loss during iterations

3.3. Explanation

The change speed of monthly average temperature anomaly seems to slow down according to the LSTM result, but never returns under zero. It is confirmed that the global warming is an irreversible process. It could only be retarded instead of being reversed. According to WMO, in 2023, the El Niño condition has occurred [15]. In fact, with the increasing of global temperature, more and more extreme climate phenomena would exist. The most well-known extreme is El Niño. Take the El Niño for example. The existence of El Niño would contribute to extreme high temperature throughout the world, and the phenomenon would last for several months. During this period, the climate and ecosystems of the planet would undergo a dramatic change as a result. According to the report published by NOAA in 2023, there is a possibility of 95% that the El Niño would last for the entire winter this year [16], which means there would be a winter-long drought. In the last hundred years, drought contribute to over 40% of the global warming [17]. The drought would cause extreme hydrological event, such as flood [18]. During the water cycle, a certain amount of CO_2 would be produced, which may have some influence on global climate change, especially on temperature. Additionally, such extremes would contribute to huge social unrest and economic loss. If the climate keeps on changing for the worst, these extreme phenomena would intensify.

This article makes use of PCA to analyze the most possible contributor of global temperature rising. CO_2 is everywhere and people emit it at any time. CO_2 is a gas that absorbs and radiates energy. According to NOAA Global Monitoring Lab, in 2021, CO_2 was two-thirds of the greenhouse gases produced by human that caused the heat. In addition, since the beginning of the first industrial revolution, the ocean surface pH has dropped with the increased amount of CO2 [3].

Moreover, this article chose the forest fire and volcanic eruption as the other two contributors of the global warming. As is shown in Fig. 3, the more forest loss due to fire, the higher global temperature is. When the forest is burning, the trees will first become carbon. Carbon will react with oxygen under high temperatures to form large amounts of carbon oxides, including carbon dioxide. And the case come back to the CO2 case. However, when a volcano is erupting, it releases a large amount of dust, which will block out the energy from the Sun. Based on the zero-dimensional energy balance theory, although the global temperature would increase for a while during the eruption, the temperature will drop rapidly since there is little sun energy coming into the planet. Thus, the more volcanic eruption, the lower temperature the Earth would have. The trend in Fig. 4 confirms the theory.

3.4. Implications

Global warming mainly caused by carbon dioxide emissions will lead to a series of change of the nature. The glacier will melt, wildfires and extreme rainfall or drought in some regions will happen more frequently. In research about snow absorption and albedo reduction, dust-snow mixing can absorb more radiation, thus enhancing the snow albedo fraction of incident solar radiation reflected to a very big extent [19]. Glaciers are an essential source of fresh water for the Earth. Once the glaciers decrease, the freshwater system will be influenced, and the sea level will increase. These experiments tell people what situation humans are facing and point out a way that all human beings have to go. Knowing that what contributes the most to global warming is CO2 emission, there are many things that can be done to reduce the emission of carbon dioxide, especially carbon dioxide emitted by human beings.

Aiming to reduce carbon emissions, there are several suggestions for different groups of people. Governments should obey and improve the supervision and management system related to carbon emissions. It is a good way to promote new energy or renewable energy transportation actively. As to some natural disasters, such as extreme rainfall and earthquakes, governments should make full emergency preparations and give timely responses. They should try their best to minimize the casualties and economic loss. For enterprise, they should abide by the policy of their governments on limiting cardon emissions, and always pay attention to the details of carbon emissions in the production activities. In addition, citizens should implement the “low- carbon” life philosophy and

lifestyle, and improve their awareness of energy conservation. They can go out with electronic cars or take more public transport.

4. Limitations and Expectations

Although all the results this article has gotten are reasonable and comprehensive, there are still some drawbacks of these experiments. First of all, the dataset one got is not as large as expected. In this article, LSTM model is used as the main model for prediction. Handling a large amount of data, such as tens of thousands or more data and still has good performance is one of the most excellent features of the LSTM model. Nevertheless, if there are not enough data, it is very likely that the results will blow up or becomes unrealistic. Additionally, this article uses temperature anomaly as the research target. The temperature anomaly is based on a baseline temperature, which is usually the average temperature of the last century. However, when forecasting the temperature anomaly in the next 100 years, the baseline temperature would change a bit. Maybe the change is very small, but it may still have certain influences on the accuracy of prediction results. Moreover, LSTM model need time to work. The more data to process, the more time it needs to work. Since LSTM model has perfect performance in deep learning, it is very easy for it to over fit the data. It is a strong model, but it also needs perfect data. When doing the PCA, there is a small amount of information loss after dimension of initial matrix is reduced. So the final data is different from the origin ones. There are lots of factors that are able to affect the climate. However, few factors have relatively well-documented data, which should have relatively large time range without large amount of data loss. Only few factors are suitable for the condition. Because of the small amount of the factors, the role and significance if the PCA method seems not so obvious.

Climate change is a big topic that involves many factors. It cannot be predicted or simulated precisely. However, that is not to say that there's nothing can be done. The prediction of future temperature, regardless its accuracy, shows that the global warming would slow down and remain in a stable stage. This result also indicates that people's contributions to slowing and curbing global warming make sense. Although it might be hard for people to solve all these drawbacks now, with the help of new technologies that would be developed, it will not be complex then.

5. Conclusion

To sum up, this study makes a prediction of temperature in the next 100 years and studies on the factors that contribute to temperature increase. Since temperature anomaly is more important than absolute temperature, all the temperature variable is temperature anomaly. This article uses LSTM model for temperature prediction. The global average temperature anomaly will get into a stable stage with the highest anomaly of 1.24℃. PCA is used for contributor analysis. The factor that influences the temperature most is CO_2. As a result, people have to develop ways to slow down global warming by reducing carbon emissions. The accuracy and reliability are large limitations of this study. But with the help of new technologies, these drawbacks would be solved then. Global warming has become a very important climate issue in recent years. These results provide some useful suggestions to different groups of people. The key to slow down global warming is to practice low-carbon life to reduce carbon emissions.

References

[1] Mikhaylov A, Moiseev N, Aleshin K, et al. Global climate change and greenhouse effect. Entrepreneurship and Sustainability Issues, 2020, 7(4): 2897.

[2] NOAA, Greenhouse gases continued to increase rapidly in 2022, published April 5, 2023, retrieved on August 10, 2023. Retrieved from: https://www.noaa.gov/news-release/greenhouse-gases-continued-to-increase-rapidly-in-2022

[3] NOAA, Rebecca Lindsey, Climate Change: Atmospheric Carbon Dioxide, published May 12, 2023, retrieved on August 10, 2023. Retrieved from: https://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide

[4] IPCC, Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)], published March 20, 2023, retrieved on August 10, 2023. Retrieved from: https://www.ipcc.ch/report/ar6/syr/

[5] IPCC. Renewable energy can mitigate global warming. Energy and Environment, 2019, 3:1 1.

[6] Swiss Department. Swiss ecologists proposed that adding 1 billion hectares of forest will effectively curb global warming. Shanghai Energy Conservation ,2019(08):718.

[7] Deng X, Jiang S, Liu Bi, Wang Z, Shao Q. Statistical analysis of the relationship between carbon emissions and temperature rise with the spatially heterogenous distribution of carbon dioxide concentration. Journal of Natural Resources, 2021, 36(4): 934-947.

[8] Gao M, Yang H, Xiao Q, et al. A novel fractional grey Riccati model for carbon emission prediction. Journal of Cleaner Production, 2021, 282: 124471.

[9] Huang Y, Shen L, Liu H. Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China. Journal of Cleaner Production, 2019, 209: 415-423.

[10] Algieri B, Iania L, Leccadito A. Looking ahead: Forecasting total energy carbon dioxide emissions. Cleaner Environmental Systems, 2023, 9: 100112.

[11] Liu Z, Wang F, Tang Z, et al. Predictions and driving factors of production-based CO2 emissions in Beijing, China. Sustainable Cities and Society, 2020, 53: 101909.

[12] Fan D, Sun H, Yao J, et al. Well production forecasting based on ARIMA-LSTM model considering manual operations. Energy, 2021, 220: 119708.

[13] NOAA, Anomalies vs. Temperature, retrieved on August 21, 2023. Retrieved from: https://www.ncei.noaa.gov/access/monitoring/dyk/anomalies-vs-temperature

[14] Hodson T O. Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. Geoscientific Model Development, 2022, 15(14): 5481-5487.

[15] World Meteorological Organization, World Meteorological Organization declares onset of El Niño conditions, published 4 July 2023, retrieved on August 21, 2023. Retrieved from: https://public.wmo.int/en/media/press-release/world-meteorological-organization-declares-onset-of-el-ni%C3%B1o-conditions

[16] NOAA, Emily Becker, August 2023 El Niño update: back to school, published August 10, 2023, retrieved on August 21, 2023. Retrieved from: https://www.climate.gov/news-features/blogs/august-2023-el-ni%C3%B1o-update-back-school

[17] Huang J, Guan X, Ji F. Enhanced cold-season warming in semi-arid regions. Atmospheric Chemistry and Physics, 2012, 12(12): 5391-5398.

[18] Chen Y, Li Y, Li Z, et al. Analysis of the impact of global climate change on arid regions. Advances in Earth Science, 2022, 37(2): 111.

[19] He C, Liou K N, Takano Y, et al. Enhanced snow absorption and albedo reduction by dust‐snow internal mixing: Modeling and parameterization. Journal of Advances in Modeling Earth Systems, 2019, 11(11): 3755-3776.